

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2004, 09:51:49 ; Search time 60 Seconds
(without alignments)
1059.553 Million cell updates/sec

Title: US-10-006-922A-12
Perfect score: 1214
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHHLFL 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	225	3 AAY99836	Aay99836 Discosoma
2	1214	100.0	225	3 AAB01622	Aab01622 Discosoma
3	1214	100.0	225	4 AAG65509	Aag65509 Anthozoan
4	1214	100.0	225	5 ABB08834	Abb08834 Yeast opt
5	1214	100.0	225	5 AAE28833	Aae28833 Discosoma
6	1214	100.0	225	5 AAE17540	Aae17540 Discosoma
7	1214	100.0	225	5 AAO18270	Aao18270 Discosoma
8	1214	100.0	225	6 AAE34962	Aae34962 Discosoma
9	1214	100.0	225	7 ADC24126	Adc24126 Discosoma
10	1214	100.0	225	7 ABW00918	Abw00918 Discosoma
11	1214	100.0	487	5 ABB08821	Abb08821 Autofluor
12	1214	100.0	506	5 ABB08822	Abb08822 Autofluor
13	1214	100.0	547	5 ABB08823	Abb08823 Autofluor
14	1211	99.8	225	5 AAE28920	Aae28920 Discosoma
15	1211	99.8	225	7 ABW00937	Abw00937 Discosoma
16	1211	99.8	225	7 ABW00929	Abw00929 Discosoma
17	1211	99.8	225	7 ABW00938	Abw00938 Discosoma
18	1210	99.7	225	5 AAE28919	Aae28919 Discosoma
19	1210	99.7	225	7 ABW00930	Abw00930 Discosoma
20	1210	99.7	226	4 AAG65510	Aag65510 Anthozoan
21	1210	99.7	242	7 ADE24109	Ade24109 Discosoma
22	1209	99.6	225	5 AAE28922	Aae28922 Discosoma
23	1209	99.6	225	7 ABW00936	Abw00936 Discosoma
24	1209	99.6	225	7 ABW00931	Abw00931 Discosoma
25	1209	99.6	226	5 ABB08835	Abb08835 Yeast opt

ALIGNMENTS

RESULT 1
AAY99836
ID AAY99836 standard; protein; 225 AA.
XX
AC AAY99836;
XX
DT 12-SEP-2003 (revised)
DT 19-SEP-2000 (first entry)
XX
DE Discosoma sp. "red" novel fluorescent protein drFP583.
XX
KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;
KW fluorescent labeling.
XX
OS Discosoma sp; "red".
XX
FH Key Location/Qualifiers
FT Misc-difference 122 /note= "encoded by TC"
FT Misc-difference 127 /note= "encoded by GTTG"
XX
PN WO200034326-A1.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US029473.
XX
PR 11-DEC-1998; 98US-00210330.
PR 14-OCT-1999; 99US-00418529.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;
PI Ding L;
XX
DR WPI; 2000-423381/36.
XX
PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,
PT useful for fluorescent labeling and as markers.
PS Claim 20; Page 74-75; 86pp; English.
XX
CC The present sequence is a novel fluorescent protein (nFP) encoded by the
CC full-length cDNA drFP583. drFP583 was isolated from Discosoma sp. "red",
CC a non-bioluminescent species of the Class Anthozoa. Fluorescent proteins
CC can be used in fluorescent labeling, a useful tool for marking a protein,
CC cell or organism of interest. Unlike other markers used in protein

26 1208 99.5 225 7 ABW00932 Abw00932 Discosoma
27 1208 99.5 225 7 ABW00935 Abw00935 Discosoma
28 1208 99.5 225 7 ABW00939 Abw00939 Discosoma
29 1207 99.4 225 5 AAE28921 Aae28921 Discosoma
30 1207 99.4 225 5 AAE17541 Aae17541 Discosoma
31 1207 99.4 225 7 ABW00941 Abw00941 Discosoma
32 1207 99.4 225 7 ABW00940 Abw00940 Discosoma
33 1206 99.3 225 7 ABW00933 Abw00933 Discosoma
34 1206 99.3 225 7 ABW00934 Abw00934 Discosoma
35 1205 99.3 240 6 ABP56678 Abp56678 Mammalian
36 1204 99.2 225 5 AAE28923 Aae28923 Discosoma
37 1204 99.2 225 5 AAE28925 Aae28925 Discosoma
38 1202 99.0 225 5 AAE28924 Aae28924 Discosoma
39 1202 99.0 240 6 ABP56685 Abp56685 Discosoma
40 1200 98.8 225 5 ABP70039 Abp70039 Colour Fa
41 1199 98.8 225 5 AAE28926 Aae28926 Discosoma
42 1199 98.8 240 6 ABP56684 Abp56684 Discosoma
43 1199 98.8 240 6 ABP56683 Abp56683 Discosoma
44 1194 98.4 240 6 ABP56681 Abp56681 Discosoma
45 1193 98.3 240 6 ABP56679 Abp56679 Discosoma

CC labeling, such as beta-galactosidase and luciferase, fluorescent proteins
CC do not require an exogenous cofactor or substrate. Methods involving
CC fluorescent proteins are also less laborious and less difficult to
CC control than the traditional methods of fluorescent labeling, where a
CC protein of interest is purified and then covalently conjugated to a
CC fluorophore derivative. Novel fluorescent proteins isolated from species
CC of the Class Anthozoa can be used as markers for gene expression and
CC protein localization studies, and in fluorescence resonance energy
CC transfer (FRET) reactions. They may have improved properties and better
CC suitability for larger excitations compared to prior art fluorescent
CC proteins such as green fluorescent protein. (Updated on 12-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.4e-128;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPRYEGHNTVKLKVTGGPLPFAWDI 60
Db 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPRYEGHNTVKLKVTGGPLPFAWDI 60
QY 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCIFY 120
Db 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCIFY 120
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 2
AAB01622
ID AAB01622 standard; protein; 225 AA.
XX
AC AAB01622;
XX
DT 12-DEC-2000 (first entry)
XX
DE Discosoma sp. red fluorescent protein drFP583.
XX
KW Anthozoa; fluorescent protein; fluorescence labeling; amFP486; cFP484;
KW zFP506; zFP538; drFP583; dsFP483; asFP600; dgFP512; dmFP592.
XX
OS Discosoma sp.
XX
PN WO200034526-A1.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US029405.
XX
PR 11-DEC-1998; 98US-00210330.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV;
XX
DR WPI; 2000-423451/36.
XX
PT Novel method for identifying a DNA sequence encoding fluorescent proteins
PT from non-bioluminescent Anthozoa which are useful for fluorescent
PT labeling and as markers.
XX
PS Claim 3; Page 68-69; 73pp; English.
XX
CC The present sequence is Discosoma sp. red fluorescent protein drFP583. It
CC was isolated using the Aequoria victoria green fluorescent protein (GFP)

Thu

CC sequence, which was used to design PCR primers which might isolate other
CC fluorescent proteins from a number of species of Anthozoa. These were
CC Anemonia majano, Clavularia sp., Zoanthus sp., Discosoma sp. and Anemonia
CC sulcata. The cDNA obtained was then screened in the search for sequences
CC encoding fluorescent proteins. The other proteins found in this manner
CC were cFP484, zFP506, zFP538, amFP486, dsFP483, asFP600, dgFP512 and
CC dmFP592. These proteins can be used as fluorescent labels (for gene
CC expression and protein localisation studies and in fluorescence resonance
CC energy transfer (FRET) studies) in place of fluorophore derivatives and
CC luciferases, as these involve laborious processes and the latter require
CC cofactors. They can also be used in place of GFP, which is too stable to
CC be useful when studying short-term or repetitive events
XX
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.4e-128;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPRYEGHNTVKLKVTGGPLPFAWDI 60
Db 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPRYEGHNTVKLKVTGGPLPFAWDI 60
QY 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCIFY 120
Db 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCIFY 120
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 3
AAG65509
ID AAG65509 standard; protein; 225 AA.
XX
AC AAG65509;
XX
DT 30-NOV-2001 (first entry)
XX
DE Anthozoa red fluorescent protein sequence.
XX
KW Fluorescent protein; Anthozoa; fluorescence; marker; FRET; red.
KW Anthozoa.
OS
XX
PN WO200162919-A1.
XX
PD 30-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004625.
XX
PR 23-FEB-2000; 2000US-0184732P.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Nelson D, Zamaira E, Tsien R;
XX
DR WPI; 2001-557704/62.
XX
PT Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise
PT functional red fluorescent proteins, and the encoding nucleic acids, with
PT key mutations for improving the proteins function.
XX
PS Disclosure; Page 85; 90pp; English.
XX
CC The invention provides a nucleic acid encoding functional red fluorescent
CC protein (II) that differs from the sequence of an Anthozoa red
CC fluorescent protein by at least one amino acid substitution, and with

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2004, 09:51:51 ; Search time 45 Seconds
(without alignments)
1577.591 Million cell updates/sec

Title: US-10-006-922A-12
Perfect score: 1214
Sequence: 1 MFSSKNVKEFMRFKVRMEG.....EDYTIQEYVERTEGRHFL 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	225	5 Q9U6Y8	Q9u6y8 discosoma s
2	1085.5	89.4	230	5 Q9GTJ7	Q9gtj7 discosoma s
3	793	65.3	221	5 Q95P04	Q95p04 goniopora t
4	729.5	60.1	232	5 Q9U6Y7	Q9u6y7 discosoma s
5	690	56.8	225	5 Q963F5	Q963f5 montastraea
6	681	56.1	225	5 Q95UA7	Q95ua7 montastraea
7	681	56.1	225	5 Q7Z0W4	Q7z0w4 montastraea
8	679.5	56.0	227	5 Q7Z0W6	Q7z0w6 montastraea
9	678.5	55.9	227	5 Q962P9	Q962p9 montastraea
10	678.5	55.9	227	5 Q7Z0W8	Q7z0w8 montastraea
11	678	55.8	225	5 Q7Z0W5	Q7z0w5 montastraea
12	669.5	55.1	234	5 Q7Z0W7	Q7z0w7 montastraea
13	668.5	55.1	234	5 Q8T5P2	Q8t5f2 montastraea
14	667	54.9	266	5 Q9U6Y3	Q9u6y3 clavularia
15	663	54.6	236	5 Q8T6U0	Q8t6u0 dendroneph
16	655.5	54.0	227	5 Q95VT0	Q95vt0 montastraea

17	654.5	53.9	234	5 Q8MU47	Q8mu47 montastraea
18	652.5	53.7	225	5 Q7Z0W9	Q7z0w9 montastraea
19	639	52.6	224	5 Q8MU48	Q8mu48 montastraea
20	631.5	52.0	225	5 Q8T5F1	Q8t5f1 montastraea
21	608.5	50.1	231	5 Q8ISF8	Q8isf8 parascyoni
22	599.5	49.4	225	5 Q8I6J8	Q8i6j8 trachyphyl
23	579.5	47.7	259	5 Q8MMA2	Q8mma2 agaricia fr
24	574	47.3	228	5 Q9GPI6	Q9gpi6 anemonia su
25	573.5	47.2	235	5 Q8T5F0	Q8t5f0 scolymia cu
26	570.5	47.0	231	5 Q8T6T8	Q8t6t8 discosoma s
27	566.5	46.7	239	5 Q8MMA1	Q8mmal agaricia ag
28	559.5	46.1	231	5 Q8T5E9	Q8t5e9 ricordea fl
29	556.5	45.8	227	5 Q95W86	Q95w86 condylactis
30	556.5	45.8	232	5 Q9GPI5	Q9gpi5 anemonia su
31	556.5	45.8	232	5 Q9GZ28	Q9gz28 anemonia su
32	553.5	45.6	227	5 Q95W85	Q95w85 radianthus
33	551.5	45.4	231	5 Q8T5E8	Q8t5e8 ricordea fl
34	548.5	45.2	227	5 Q95W11	Q95w11 condylactis
35	546.5	45.0	231	5 Q8MU46	Q8mu46 ricordea fl
36	530.5	43.7	228	5 Q86LV4	Q86lv4 radianthus
37	530.5	43.7	234	5 Q8T5F3	Q8t5f3 scolymia cu
38	529.5	43.6	227	5 Q8MU45	Q8mu45 condylactis
39	522.5	43.0	225	5 Q8T6T9	Q8t6t9 radianthus
40	511.5	42.1	229	5 Q8T5E7	Q8t5e7 condylactis
41	495	40.8	229	5 Q9U6Y6	Q9u6y6 anemonia ma
42	490	40.4	214	5 Q86LV8	Q86lv8 meandrina m
43	490	40.4	214	5 Q86LV7	Q86lv7 meandrina m
44	485.5	40.0	231	5 Q9U6Y4	Q9u6y4 zoanthus sp
45	485.5	40.0	238	5 Q9BLY9	Q9bly9 renilla mue

ALIGNMENTS

RESULT 1
Q9U6Y8
ID Q9U6Y8 PRELIMINARY; PRT; 225 AA.
AC Q9U6Y8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fluorescent protein FP583.
OS Discosoma sp.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
OX NCBI_TaxID=86600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99436614; PubMed=10504696;
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,
RA Markelov M.L., Lukyanov S.A.;
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";
RL Nat. Biotechnol. 17:969-973(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,
RA Markelov M.L., Lukyanov S.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF168419; AAF03369.1; -.
DR PDB; 1G7K; 07-NOV-01.
DR PDB; 1GGX; 06-DEC-00.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 225 AA; 25931 MW; FBF9A5369778F689 CRC64;

Query Match 100.0%; Score 1214; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.4e-101;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60
Db 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60
QY 61 LSPQFQYGSKVYVHPADIPDYKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120
Db 61 LSPQFQYGSKVYVHPADIPDYKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120
QY 121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
Db 121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
QY 181 YMAKCPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db 181 YMAKCPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 2
Q9GTJ7 PRELIMINARY; PRT; 230 AA.
AC Q9GTJ7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Red fluorescent protein.
GN FP593.
OS Discosoma sp. SSAL-2000.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
OX NCBI_TaxID=137428;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20434599; PubMed=10981720;
RA Fradkov A.F., Chen Y., Ding L., Barsova E.V., Matz M.V.,
RA Lukyanov S.A.;
RT "Novel fluorescent protein from Discosoma coral and its mutants
possesses a unique far-red fluorescence.";
RL FEBS Lett. 479:127-130(2000).
DR EMBL; AF272711; AAG16224.1; -.
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 230 AA; 26370 MW; 5215B1B436D67E51 CRC64;

Query Match 89.4%; Score 1085.5; DB 5; Length 230;
Best Local Similarity 88.1%; Pred. No. 5.8e-90;
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60
Db 1 MSCSKNVKEFMRFKVRMEGTGNGHEFEIKGEGERPEYEGHCSVKLMVTKGGPLPFAFDI 60
QY 61 LSPQFQYGSKVYVHPADIPDYKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120
Db 61 LSPQFQYGSKVYVHPADIPDYKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120
QY 121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
Db 121 EVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
QY 181 YMAKCPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db 181 YMAKCPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 3
Q95P04 PRELIMINARY; PRT; 221 AA.
ID Q95P04
Q95P04

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE GFP-like chromoprotein.
OS Goniopora tenuidens.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Fungiina; Poritidae; Goniopora.
OX NCBI_TaxID=75301;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21538626; PubMed=11682051;
RA Gurskaya N.G., Fradkov A.F., Terakikh A., Matz M.V., Labas Y.A.,
RA Martynov V.I., Yanushevich Y.G., Lukyanov K.A., Lukyanov S.A.;
RT "GFP-like chromoproteins as a source of far-red fluorescent
proteins(1).";
RL FEBS Lett. 507:16-20(2001).
DR EMBL; AF383156; AAL27542.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 221 AA; 24918 MW; 93F9F4B5C2003CB4 CRC64;
Query Match 65.3%; Score 793; DB 5; Length 221;
Best Local Similarity 68.6%; Pred. No. 1.4e-63;
Matches 144; Conservative 25; Mismatches 41; Indels 0; Gaps 0;
QY 6 NVIKEFMRFKVRMEGTGNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDILSPQF 65
Db 2 SVIAKQMTYKVYMSGTVNGHYFEVQDGKPKYEGEQTVKLVTKGGPLPFAWDILSPQS 61
QY 66 QYGSKVYVHPADIPDYKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFYKVKFI 125
Db 62 QYGSIPFTKYPEDIPDYVQSPFEGYTWERIMNFEDGAVCTVSDSSIQNCFIYNVKS 121
QY 126 GVNFPDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSIYMAKX 185
Db 122 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALEGGHYLCEFKSTYKAKK 181
QY 186 PVQLPGYVYVDSKLDITSHNEDYTIIVEQY 215
Db 182 PVKMPGYHYVDRKLDVTNHNIDYTSVEQCE 211

RESULT 4
Q9U6Y7 PRELIMINARY; PRT; 232 AA.
ID Q9U6Y7
AC Q9U6Y7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fluorescent protein FP483.
OS Discosoma striata.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
OX NCBI_TaxID=105400;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99436614; PubMed=10504696;
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaratisky A.G.,
RA Markelov M.L., Lukyanov S.A.;
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";
RL Nat. Biotechnol. 17:969-973(1999).
DR EMBL; AF168420; AAF03370.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 232 AA; 26435 MW; AA8F18EEE283CE4D CRC64;